First Hit Fwd Refs

☑ Generate Collection Print

L4: Entry 1 of 5 File: USPT Jul 15, 2003

DOCUMENT-IDENTIFIER: US 6594786 B1
TITLE: Fault tolerant high availability meter

Abstract Text (1):

A fault tolerant availability meter includes agents for stand-alone computers and each node of a cluster. The agents monitor availability with timestamps and report uptime and downtime events to a server. Additionally, agents on nodes of a cluster monitor cluster, node and package availability and cluster configuration changes and report these event to the server. Events are stored locally on the stand-alone computers and nodes, and additionally, on the server. Events are tracked with a sequence numbers. If the server receives an out-of-sequence event, an agent-server recovery procedure is initiated to restore the missing events from either the agents or the server. The server may generate availability reports for all monitored entities, including one or more stand-alone computers and one or more clusters of computers. Availability is distinguished by planned and unplanned downtime. Furthermore, unavailable and unreachable systems are identified.

Brief Summary Text (11):

First, EMFs generally do not distinguish between "unavailable" and "unreachable" systems. An EMF will treat a system that is unreachable due to a network problem equivalent to a system that is down. While this is appropriate for speedy problem detection, it is not sufficient to determine availability with any degree of accuracy. Second, because EMFs poll monitored systems over a network, their resolution is insufficient for mission critical environments. The polling intervals are usually chosen to be short enough to give prompt problem detection, but long enough to avoid saturating the local network. Polling intervals in excess of ten minutes are typical. This implies that each downtime event has a 10-minute margin of error. High availability systems often have downtime goals of less than 5 minutes per year. Thus, systems based on polling are inherently deficient to measure availability for high availability systems with a sufficient degree of accuracy. Third, while EMFs can monitor the availability of system and network resources to a certain degree, they do not have a mechanism for monitoring redundant hardware resources such as clusters, or of detecting the downtime associated with application switchover from one system to another. For example, the availability of service for a cluster may be 100% even though one of its nodes has failed. Finally, EMFs tend to be very complex, resource intensive and difficult to deploy.

Brief Summary Text (17):

According to a preferred embodiment of the present invention, a fault tolerant method of monitoring one or more computers for availability may include generating an event when a computer system detects a change in its status that affects availability; transmitting the event from the computer system to a central repository; and periodically re-transmitting the event if a receipt confirmation message is not received from the central repository. The computer system may store the event in a local repository located on the computer system before transmitting the event to the central repository. If a receipt confirmation message is not received from the central repository, the event is held in a queue for re-transmission at a later time. If the computer system receives a status request from the central repository, in addition to reporting status, the computer system will transmit the events held in the queue.

Brief Summary Text (18):
The present invention also includes a fault tolerant method of monitoring one or more computers for availability, where the method may include generating an event containing a sequence number when a computer system detects a change in its status that effects availability; transmitting the event from the computer system to a central repository; comparing the sequence number of the event with a next expected sequence number computed from reading the central repository; and synchronizing events between the computer system and the central repository if the sequence number does not match the next expected sequence number. A copy of each event may be maintained in a local repository on the computer system. If the sequence number matches the next expected sequence, the events and sequence numbers are stored in the central repository. If the sequence number is greater than the next expected sequence number, the central repository requests the missing events from the computer system. If the sequence number is less than the next expected sequence number, the central repository determines whether the event has already been received. If the event has already been received, the event is discarded. If the event has not already been received, the computer system has lost events and the central repository sends the missing events to the computer system.

Detailed Description Text (10):
Each HA agent 20 monitors the availability of the system on which it is installed and generates events when changes in system availability or configuration are detected. Additionally, if the system is a node of a cluster, the HA agent 20 also generates events when availability or configuration changes to the cluster, nodes, or packages are detected.

Detailed Description Text (12):
The HA meter M is beneficially designed to minimize utilization of network resources. Essentially, no network traffic (i.e., event) is generated between the agents 20 and the server 22 unless the state of a monitored entity changes. The HA server 22 rarely polls the HA agent 20. During normal operations, each HA agent 20 maintains availability data locally. Unless a system crash or cluster change event is detected, no events are generated.

Detailed Description Text (21):
The HA agent daemon 30 registers with the event monitoring service 40 upon installation to receive changes in cluster availability, such as clusters to which the monitored system belongs, and packages about which the monitored system has information.

Detailed Description Text (22):
Whether HA agent 20 is monitoring a stand-alone server 2 or cluster C, events are captured and reported to the HA server 22. An event may either indicate availability or unavailability of a particular resource (i.e. availability event) or may indicate a change to the configuration of a system, node, package or cluster (i.e., configuration event). Thus, availability and configuration events are related to a particular system, node, package or cluster. For example, if a node fails several events may be generated: a node event may be generated to indicate that the node has failed; a first package event may be generated indicating that a package on the failed node has failed; a second package event may be generated indicating that the package has been restarted on a backup node; and one or more configuration events may be generated indicating the change in cluster configuration. Table II indicates the types of data included in an event.

Detailed Description Text (26):
A cluster monitor 44a is a special type of monitor 44 designed to monitor cluster, node and package status on the cluster C. In the preferred embodiment, the cluster monitor 44a makes use of cluster MIBs (management information base) that are

managed by daemons responsible for tracking status of nodes and packages. Generally, the node status is tracked with a heartbeat mechanism. Package status is tracked by the process or task ID assigned by the OS when the package is launched. Status changes to the cluster, node and package are reported in the form of an event to the EMS 40, which in turn provides the information onto the interested HA agent 20.

Detailed Description Text (32):
The configuration database 64 tracks cluster configuration changes as they are received from the HA agents 20. Clusters are also comprised of packages, which loosely represent the business services (e.g., application or database) performed by the datacenter D. Each package is either running on a node or not allocated. A package is considered unavailable if it is not allocated. A cluster configuration may undergo changes and these changes are also tracked in the configuration database 64. For example, if a node has been added or removed from the cluster C. However, if a node (e.g., 4 and 6) has been added to the cluster C but no cluster monitor 44a has been configured for the node, then its state is maintained as unmonitored and it does not contribute to any availability computation. Thus, the cluster, node and package availability events interpreted in view of the configuration database 64 provides the HA meter M with an accurate view of cluster, node and package availability.

Detailed Description Text (39):
The cluster, node and package entities have different downtime definitions. Cluster downtime may be defined as any interval in which the cluster is halted or all cluster nodes are down or halted. Node downtime may be defined as any interval in which the node status (relative to a given cluster) reported by the cluster monitor 44a is not "running," or any interval in which the system on which the node is running is known to be unavailable. Package downtime may be defined as the time between package failure or halt on a given cluster node, and its startup on another (or the same) node in the same cluster. Package failure or halt is indicated when the package status reported by the cluster monitor 44a changes from "up" to "not up" on a given system, or when a system downtime event is received for the system on which the package is running. Package restart is indicated when an cluster monitor 44a sends a package notification event with status of "up." Package failover and restart time are counted as "downtime."

Detailed Description Text (41):
Availability and configuration events are sent from a monitored entity to indicate that the state of the monitored entity has changed. When an availability or configuration event is received, the HA server 22 first checks the event sequence number and initiates any necessary recovery protocols. The event is then archived, the configuration database 64 is updated, and any secondary events are generated.

Detailed Description Text (59):
FIG. 7D illustrates a flowchart of a procedure performed by the HA agent 20 in response to a cluster event. In step 174, a cluster event is received from the cluster event monitor 44a. If the cluster event monitor 44a is not designed to send event messages when status changes, alternatively, the HA agent 20 could poll the cluster event monitor 44a for status changes. At step 176, the HA agent 20 prepares to package the cluster event in an availability or configuration event in accordance with the data types listed in Table II. In particular, a sequence number is generated for the event. At step 178, the status file 36 is updated with a current timestamp and the event is logged in the event log 32. At step 180, the event is transmitted to the HA server 22.

Detailed Description Text (64):
Availability and configuration events are received from monitored systems (i.e., systems 2, 4 and 6) to indicate that the state of a monitored entity has changed, as shown in step 200. When an event is received, the HA server 22 checks the event

sequence of the received event message.

Detailed Description Paragraph Table (3):
TABLE III 1) Monitor start events a) System - The HA Agent has started monitoring this system. b) Cluster - The HA Agent has started monitoring this cluster. c) Node - The HA Agent has started monitoring this node. d) Package - The HA Agent has started monitoring this package. 2) Monitor restart events a) System - The system rebooted. b) Cluster - The cluster is back up. c) Node - The node is back up. d) Package - The package is back up. 3) Monitor shutdown events a) Planned - The system underwent an orderly shutdown. b) Unplanned - The system crashed. 4) Cluster state change events a) Cluster i) Up - The cluster is back up. ii) NotUp - The cluster went down. b) Node i) Up - The node is back up. ii) NotUp - The node went down. c) Package i) Up - The package is back up. ii) NotUp - The package went down. d) Configuration Change i) System - The system configuration changed (in some way that affects availability measurement). ii) Cluster - The cluster configuration changed.

CLAIMS:

1. A fault tolerant method of monitoring one or more computers for availability, comprising: generating an event when a computer system detects a change in its status that affects availability; transmitting the event from the computer system to a central repository; and periodically re-transmitting the event if a receipt confirmation message is not received from the central repository.

8. The method of claim 1, wherein the change of status includes changes in availability and configuration.

9. The method of claim 1, wherein an event indicating a change in availability includes a timestamp, event type and source designator.

10. A fault tolerant method of monitoring one or more computers for availability, comprising: generating an event containing a sequence number when a computer system detects a change in its status that effects availability; transmitting the event from the computer system to a central repository; comparing the sequence number of the event with a next expected sequence number computed from reading the central repository; and synchronizing events between the computer system and the central repository if the sequence number does not match the next expected sequence number.